2030 Census Research Project Explorer

2030 Census research projects are grouped into five focus areas that the Census Bureau refers to as Enhancement Areas (EAs).

These EAs guide the development of the 2030 Census research agenda.

The 2030 Census Research Project Explorer is an online tool that can be used to learn more about the research projects being conducted in each EA and explore topics of interest. It can be accessed here: https://www.census.gov/data/data-tools/2030-census-research-explorer.html

This document provides a downloadable copy of the information available in the 2030 Census Research Project Explorer online tool.

United States® Census Bureau

U.S. Department of Commerce
U.S. CENSUS BUREAU
census.gov

Revised: July 27, 2023

Table of Contents

What are Enhancement Areas? 3
EA 1: Data Collection -- Self-Response 5
EA 1: Data Collection -- In-Field and In-Office Enumeration 13
EA 1: Data Collection -- Response Data Quality and Questionnaire Content 21
EA 1: Data Collection -- Post-Enumeration Survey (PES) 28
EA 2: Modernize Group Quarters Enumeration 31
EA 3: Integrate Data Processing with Data Collection in Real Time 40
EA 4: Streamline Operational Support Infrastructure 51
EA 5: Continuous Data Collection and Aggregation 58

What are Enhancement Areas?

Enhancement Areas (EAs) refer to the five focus areas that guide the development of research used to inform the 2030 Census operational design. They also serve to keep research focused on specific outcomes, while providing transparency into program research priorities and desired outcomes.

Enhancement Area 1: Data Collection
This EA focuses on enhancing data collection methods and operations, building on 2020 Census innovations, and further improving data quality. It will research the best contact methods and data collection strategies, with an emphasis on improving data collection for specific population groups, particularly historically undercounted populations. Research conducted through this EA also aims to enhance our engagement and outreach to communities and historically undercounted populations. Several research projects focus on increasing automation in our data collection operations to reduce respondent burden and improve efficiency.

Note that this Enhancement Area comprises sub-EAs, each of which has designated research projects. These sub-EAs are:

- Self-Response,
- In-Field and In-Office Enumeration,
- Response Data Quality and Questionnaire Content,
- Post-Enumeration Survey (PES).

Enhancement Area 2: Modernize Group Quarters Enumeration
This EA focuses on group quarters (GQs), which are places where people live or stay in a group living arrangement, such as college or university student housing or a nursing home. While GQ residents represent a small proportion of the population, they are a unique population that must be handled with specialized methods and operations. This EA will build a good inventory of GQ facilities to ensure an accurate enumeration of the residents. It will also research ways to improve data collection methods for GQs to improve data quality and reduce the burden on GQ facilities and residents. Through these efforts, this EA also seeks to build partnerships with the GQ community and their advocates.

Enhancement Area 3: Integrate Data Collection and Processing in Real Time
This EA is focused on integrating data collection and data processing, improving data quality in near real time, and pivoting away from the historical linear approach for processing census data toward a more transaction-based framework. By shifting to processing data in near real time we can minimize the need for a large-scale post-collection processing operation. Additionally, by integrating analytics and near real time monitoring of data, we can identify and address quality issues and make data-centric decisions to resolve anomalies as they surface. This research will streamline data processing activities and enhance progress reporting throughout the collection process.

Enhancement Area 4: Streamline Operational Support Infrastructure
This EA is focused on transforming and improving the support infrastructure for the 2030 Census. It focuses on topics such as reducing the physical footprint for infrastructure; re-envisioning the office structure and the use of automation for field operations and other respondent services; researching
(continued on next page)

(continued)
ways to streamline and automate recruiting, hiring, and payroll services to increase efficiency and facilitate large scale staffing for peak operations; and researching ways to improve external engagement and leverage partnership networks.

Enhancement Area 5: Continuous Data Collection and Aggregation
This EA focuses on gathering and assessing data throughout the decade to optimize the once-a-decade count. Building on the 2020 Census success of conducting the majority of address frame development across the decade, this EA will research ways to apply a similar strategy to person data collection, decreasing the amount of work necessary at the end of the decade. This research will identify and procure new administrative records data to increase coverage of all populations. This EA will also establish the ability to continuously develop and assess population estimates from administrative data throughout the decade, to identify the coverage and quality gaps.

Project: Mobile Questionnaire Assistance Scope Determination

This project's objective is to determine the optimal design for the Mobile Questionnaire Assistance (MQA) operation such that it provides benefit to the 2030 Census in terms of contributing to response and providing support for communities, including historically undercounted populations, through a physical presence in the communities. MQA includes functional activities in outreach, communications, and partnership that cover awareness and promotion, respondent motivation, and assistance. This project looks to determine the design of MQA through research methods such as focus groups and data modeling. The project will aid in determining the recommended MQA design to be tested in a subsequent field test.

Research Topics

- Design of MQA operation, optimized for 2030 Census self-response and supporting communities through a physical presence.
- Mobile application to enable self-response and interview-based response.

Research Questions

1. What criteria should be used to determine MQA locations?
2. How best can we operationalize the MQA location selection process?
3. How should the criteria for determining the areas for MQA be adapted once in-field enumeration has started, including determining whether MQA should be available at all once in-field enumeration starts?
4. If MQA is running during in-field enumeration, then how can in-field enumeration completion rates be incorporated into the criteria for selecting the MQA areas? Does MQA become static during in-field enumeration operations?
5. What is the best combination of mobile and static locations to increase response and provide community support (all mobile, some mobile and some static, etc.)? What are the criteria for determining static locations and what does it mean to be static?
6. What type of non-monetary incentives (promotional items) to use at events that will encourage participation at MQA and will continue to share the message with others after leaving an MQA event?
7. How important is it for respondents to have the opportunity to respond to the census at in-person events? Are individuals inclined to respond to the census at in-person events?

Project: Improving Internet Self-Response and Non-ID Data Collection

This project's objective is to improve the internet self-response data collection instrument, expected to be the primary way of collecting responses to the 2030 Census. This project will research ways to improve the address collection for the ISR Non-ID option, as well as conduct research on a whole-or-partial household question, the Usual Home Elsewhere (UHE) address collection, the quality and completeness of coverage data, and respondent self-identification of duplicate, vacant, or nonexistent housing units.

Research Topics

- Improve the Non-ID user experience for respondents by simplifying the address collection and verification screens to improve the quality of addresses collected through the Internet Non-Response Non-ID option.
- Redesign a portion of the application to improve vacancy reporting by providing a user-friendly way for Non-ID respondents to report vacancy status.
- Provide a confirmation screen for respondents reporting for vacant units.
- Improve the identification of potential coverage issues including count discrepancies, undercount, and overcount, by enhancing the coverage questions and count discrepancy checks, collecting addresses for respondents who indicate they usually live or stay somewhere else, and ensuring sufficient information is provided for geocoding.
- Design and test question(s) that allow respondents to self-identify an address that should be considered for deletion.
- Design and test the use of a question to indicate whole or partial household response.
- Usual Home Elsewhere (UHE) address collection.

Research Questions

1. Will the design enhancements improve the respondent experience, reduce respondent burden, and increase the quality of the data provided by the respondent?
2. How could respondents self-report units that no longer exist and should be considered for deletion?
3. What types of coverage issues can be identified during the initial census response? Can any be resolved during the initial census response?
4. What information needs to be collected during the initial census response to resolve the different coverage issues?

Project: Paper Questionnaire Fulfillment Strategy

This project's objective is to examine the feasibility of implementing a fulfillment operation for the 2030 Census that would provide a paper questionnaire upon request from the public. A paper questionnaire fulfillment function would allow respondents to request a paper questionnaire be mailed to their housing unit, potentially via an Interactive Voice Response (IVR) system or by interacting with Census Questionnaire Assistance (CQA) Customer Service Representatives.

Research Topics

- Develop a strategy for distributing paper questionnaires to respondents upon request, to include proposed methods for collecting requests from housing units and mailing out the requested questionnaires.
- Explore system solutions that could be employed to collect requests.
- Propose timeframes during which paper questionnaire fulfillment would be available to the public.

Research Questions

Respondent Requests

1. What was the demand for paper questionnaire fulfillment during the 2020 Census? How many requests were made and by what means (e.g., through CQA or other methods)? How many respondents in the Internet First cohorts returned paper questionnaires after the 4th and 7th mailings? What is the anticipated demand for paper questionnaire fulfillment during the 2030 Census?
2. Would the public-facing questionnaire request system be primarily self-service (e.g., through the CQA IVR system), or would a respondent need to speak directly with a call center agent to request a form?
3. How might an option to request a paper questionnaire be communicated to the public as part of the overarching communication campaign?
4. What are the requirements for a respondent to be eligible to request a paper questionnaire?

Mail Strategy

5. How can a questionnaire fulfillment operation be best structured to provide paper questionnaires upon request in a timely manner?
6. How would a fulfillment operation fit into the overall self-response contact strategy?
7. At what point(s) during the data collection time period should a paper questionnaire be available for request, and how would the ability to request a questionnaire be communicated as part of the self-response mail strategy?
8. Should a paper questionnaire fulfillment offering impact the timing or universe for a blanket questionnaire mailing (i.e., the questionnaire sent as part of Mailing 4 during the 2020 Census)?
9. How would a paper questionnaire fulfillment function impact printing and mailing costs?

Systems Considerations

10. Which systems are potentially impacted by paper questionnaire fulfillment?
11. What request scenarios need to be considered, and what are the systems paths for each (e.g., request with Census ID and address respondent-confirmed; request with Census ID and respondent-recommended address update; request without Census ID for known address; request without Census ID for unknown address; etc.)?

(continued on next page)

(continued)

Project: Paper Questionnaire Fulfillment Strategy

12. How is the address information collected from a respondent and verified as accurate? Does the requestor need to live at the address? How will it be confirmed if a response has not been received?
13. What happens when a respondent requests a paper questionnaire for an address that is not in the census address list? How is that request and subsequent response handled?
14. How is a non-response from a housing unit that has requested a paper questionnaire handled? If the address was previously not known, does that new address get sent to in-field enumeration for follow-up if a response is not received?
15. How much might the non-ID workload be impacted by a questionnaire fulfillment operation?
16. What is the proposed cadence of transmitting paper request workloads to the mailing operation?
17. What is the proposed mailout cadence?

Project: Provide Response Status Across Modes

This project's objective is to explore providing a mechanism whereby a respondent could confirm the Census Bureau has received a response to the census from their housing unit. The type of response status that the functionality would provide would be dependent on privacy and security limitations, and data available from other integrated systems. The modes of response status provision and recommended availability (e.g., directly to the respondent or by speaking with staff from a census operation) will also be explored. Making this functionality available is expected to improve the user experience and reduce respondent burden.

Research Topics

- Response status provision functionality.
- Policy and security positions.
- System considerations.
- Operational impacts.

Research Questions

1. What are the benefits of providing a response status to respondents? (Reduced burden, improved respondent experience, reduced operational cost, increased efficiency?)
2. What methods can be used to provide response status to respondents while still protecting privacy?
3. How would a respondent need to identify themselves before receiving information about their response status? How many and what types of identifiers should be used?
4. What information about a housing unit's response status would be available via a response status request?

Project: Machine Learning Using Census Questionnaire Assistance Audio Files

This project's objective is to use natural language processing (NLP) and other machine learning tools to analyze Census Questionnaire Assistance (CQA) operational information with the goal of using the results to suggest improvements for the CQA operation and enhance the overall call experience for callers and CQA Customer Service Representatives (CSRs).

Research Topics

- Improve fully automated or partially automated processes, such as the Interactive Voice Response (IVR) system and Frequently Asked Questions (FAQs).
- Automate some manual processes, such as the flagging system to alert a supervisor to a threatening call.
- Identify caller concerns and improve CSR scripts.

Research Questions

1. Can machine learning models be developed that allow call topics to be accurately identified in English and Spanish, and in other non-English languages, as time permits?
2. Can machine learning models be used to identify call types (e.g., threatening calls, enumerations) in real time? Can machine learning models be used to identify deviations from prescribed text in real time?
3. Can the project team use machine learning and other methods to identify caller concerns, whether these concerns were satisfied by the scripts CSRs read aloud, and whether the concerns differed by language, time, or region? Do call topics map to existing FAQs and FAQ keywords? Can FAQs be improved based on topic modeling?

Project: Optimize Census Mail Strategies

This project's objective is to determine recommended 2030 Census Mail Strategies, including improvements to optimize contact and elicit self-response. The project will explore the use of an Every Door Direct Mailer (EDDM) for the Self-Response and Update Leave operations of the 2030 Census, and conduct retrospective analysis of the 2020 Census mailing cohorts, the mailing materials language flag (commonly referred to as the bilingual flag), and the 2020 Census contact strategy assignments (Internet First and Internet Choice).

Research Topics

- Optimal self-response contact strategy.
- Improvements to specific contact strategy elements.
- Performance of 2020 Census contact strategy relative to expectations.

Research Questions

1. What is the optimal contact strategy to elicit response in self-response areas?
2. Should any additions, deletions, or adjustments be made to the 2020 Census contact strategy to improve response, reduce cost, reduce burden, or better fit expected response patterns in 2030?
3. How did the 2020 Census contact strategy perform relative to expectations (e.g., were response rates as expected? did they differ in unexpected ways by panel or cohort?)?
4. How could an EDDM be used as part of the mail strategy?
5. Did staggered mailings in 2020 effectively limit spikes in the number of calls to the Census Questionnaire Assistance operation?
6. Were there contact strategy–response mismatches?
    - Were there areas that responded using Spanish instruments at relatively high rates that did not receive bilingual materials?
    - Were there areas that responded using Spanish instruments at relatively low rates that did receive bilingual materials?
    - Were there areas that responded using the paper questionnaire at relatively high rates in the Internet First contact strategy?
    - Were there areas that responded using the paper questionnaire at relatively low rates in the Internet Choice contact strategy?
    - Would an EDDM reach households that may not have been included in the census address lists or assigned to the appropriate Type of Enumeration Area (TEA), including areas with high rates of P.O. Boxes?
7. What impact might an EDDM have on non-ID response rates?
8. Could an EDDM be implemented for a subset of areas in a dynamic manner as determined by metrics during data collection (e.g., areas with high Undeliverable As Addressed [UAA] rates, areas with low response rates)?
9. Is the design enhancement of the EDDM (e.g., glossy, colorful) expected to increase self-response rates compared to the typical postcards?
10. Is there prior research that indicates an EDDM mail piece may be more effective than a standard postcard for eliciting a response?
11. How might utilizing EDDM impact the cost and operational complexity of the print/mail program?

Project: Tailored Contact Strategies

This project's objective is to examine the use of tailored contact strategies – that is, messaging, materials, and methods of inviting households to self-respond to the census designed with the needs and preferences of different demographic groups and geographic areas in mind – that incorporate in-office enumeration methods and estimated response propensities. As part of this project, a plan to have a contact strategy for individual housing units that depends on the characteristics of that housing unit rather than a relying on a strategy applied at the national level may be developed.

Research Topics

- Data sources and data quality.
- Messaging, methods, and contact strategies.

Research Questions

1. How does changing quality cutoffs on records in a composite administrative records file affect final response quality, especially data missingness of the administrative data combined with the collected data?
2. What is the relationship between administrative records quality and a likelihood of responding score? How can both measures be used to target messaging and methods?
3. Can messaging, methods (which would include sending cases directly to field data collection based on response propensity), and number of attempts be targeted to specific populations including, but not limited to, bilingual households, those missed in administrative records, those with a lower propensity to self-respond?
4. Can self-response and enumerator modes be blended to develop a comprehensive contact strategy that is able to be adjusted during data collection to encourage response?
5. Is the use of administrative records to target messaging and methods for populations expected to increase the population's response rate?
6. Is the use of a likelihood of self-responding score to determine number of contacts for a population expected to affect a population's response rate?

Project: Full Listing and Enumeration Integration (FLEXI) for Field Data Collection Operations

The automation of listing and enumeration for the 2020 Census represents a significant advancement in technology for decennial operations. To further these advancements, the Full Listing and Enumeration Integration (FLEXI) project was created to analyze the potential of an integrated approach to listing and enumeration operations for the 2030 Census. FLEXI will focus on field housing unit data collection, but research from this project may benefit the group quarters (GQ) operations and Island Area Censuses. This project will not direct how operations are conducted but will coordinate with the operations to investigate how an integrated system and device would fit within their requirements.

Research Topics

- Benefits of integrated digital listing and enumeration capabilities.
- Data output from field operations that are necessary for real-time processing.
- Additional information useful to housing unit data collection (at the time of enumeration).
- Whether integrated listing and enumeration capabilities as well as related improvements (such as real-time or near real-time processing) change the ideal timeline for 2030 Census field operations.

Research Questions

1. From a system's perspective, how can digital listing and enumeration capabilities be integrated?
2. From a user's perspective, how can digital listing and enumeration capabilities be integrated?
3. How can field staff be trained on both listing and enumeration in the same training? What are the effects and challenges?
4. What output from field housing unit data collection meets the needs for near real-time processing and data analytics?
5. What additional information would help listers and enumerators complete their assignments?
6. Can one mapping solution fulfill the needs for field data collection?
7. Are there changes from the historical timeline of field data collection with integrated listing and enumeration functionality?

Project: Island Areas Censuses System for Enumeration and Listing

This project will explore the feasibility of using automated address listing and response collection instruments for the 2030 Island Areas Censuses (IAC) operation, including leveraging existing automated American Community Survey or decennial census address listing and collection instruments for this purpose.

Research Topics

- Review the 2020 IAC Assessment and Lessons Learned report.
- Assess the feasibility of the potential solutions for automation of data collection.
- Develop requirements for automated listing and collection instruments for IAC.

Research Questions

1. Is there a feasible way to implement automated listing and collection instruments for use in the Island Areas?
2. Can decennial census or American Community Survey instruments and systems be leveraged?
3. Are there concerns regarding the existing network infrastructure that would present challenges?
4. Are there unique IT security concerns?
5. What logistical challenges would be expected?

Project: Administrative Data Enumeration Applications

Decennial census efforts to count the population rely primarily on methods to encourage people to self-respond, followed by in-person interviewing for nonresponding households. To increase efficiency and reduce costs, the 2020 Census leveraged administrative data to enhance respondent-provided information, validate respondent submissions, and to reduce in-person field followup visits. The 2030 Census will continue these 2020 Census innovations while also exploring additional innovations to further streamline census data collection and minimize respondent burden.

Research Topics

- Quality of administrative data compared to data collected from household or proxy respondents.
- Number of attempts needed to resolve cases in the In-Field Enumeration operation's workload.
- Feasibility of using administrative and operational data to resolve cases in the In-Field Enumeration workload in lieu of additional field attempts.
- Ways to reduce respondent burden using administrative data.

Research Questions

1. What is the effect on data quality of using administrative data in lieu of In-Field Enumeration household or proxy interview data?
2. Can the number of contact attempts for In-Field Enumeration cases be reduced by using operational data in conjunction with administrative data to complete cases?
3. Can the questions asked during an In-Field Enumeration interview be reduced if high-quality administrative data exist for the housing unit?
4. Can information from administrative records be incorporated into the In-Field Enumeration interview?
5. Investigate other potential uses of administrative data to improve In-Field Enumeration interviews, reduce the number of contact attempts, and increase the efficiency of In-Field Enumeration related clerical work.

EA 1: Data Collection In Field and In Office Enumeration

Project: Case Assignment and Contact Strategy

This project aims to improve case assignment methodologies and contact strategies for the 2030 In-Field Enumeration operation. The 2020 NRFU operation was conducted using automated case management, case assignment, and data collection. While successful, the implementation was hampered by development and cultural constraints. The motivation for this project is to use the lessons learned from the 2020 NRFU operation and the results of post-2020 research to improve the efficiency of contact strategies and case management for the 2030 In-Field Enumeration operation.

Research Topics

- Identify strengths and weaknesses of the 2020 NRFU case assignment rules and contact strategy.
- Improve the efficiency of case assignments, including updates to assignments in multiunit housing.
- Visualize remaining workload on maps in a meaningful way to field staff.
- Integrate the new case assignment methodology into the field staff's workflow.
- Investigate which field staff should be able to make manual assignments.
- Effectively incorporate the new case assignment methodology into training materials.
- Improve and implement the integrated contact strategy for the 2030 Census.

Research Questions

1. What are the operational efficiencies of changing the geography of case assignment from individual housing units (HUs) to small groupings of HUs in the case assignment methodology?
2. What changes should be made to improve the process by which cases are manually assigned?
    - What are constraints for manual assignment of cases based on small groupings from question 1?
    - What are the benefits and deficiencies in terms of efficiency and quality control in changing or expanding who can manually assign cases in the ACO?
3. What messaging and content are needed to communicate changes in case assignment methodology and how it should be used by field staff?
4. Multiple research proposals are addressing aspects of the In-Field Enumeration collection strategy. Given the planned research for adaptive design improvements to data collection, what is the feasibility of implementing them in an iterative manner as they are improved versus all of the changes at the same time?
5. What is the most efficient method to implement improvements to the In-Field Enumeration collection strategy?

EA 1: Data Collection In Field and In Office Enumeration

Project: Universe Conversions During Enumeration Operations

During the 2020 Census, the type of living quarter (LQ) determined whether it was visited during the Nonresponse Followup (NRFU), Group Quarters Enumeration (GQE), Update Enumerate (UE), Update Leave (UL), or the Enumeration of Transitory Locations (ETL) operations. There was limited ability to transfer living quarters that were found in the incorrect universe because of the design of those operations. The inability to correct all misclassified LQs led to: (1) confusing operational procedures for field staff, (2) reduced quality and completeness of enumerations, and (3) additional respondent burden from repeated, uncoordinated contact attempts.

Research Topics

- Identify the types of living quarters that were eligible for each field operation in the 2020 Census.
- Evaluate the scope of universe conversions made and those that could not be made in the 2020 Census.
- Develop recommendations for training field staff to identify the type of living quarter while in the field and the conversion process when a living quarter has been misclassified.

Research Questions

1. How often were housing units, transitory locations, and group quarters misclassified during the 2020 Decennial Census?
2. How many misclassified LQs resulted in a conversion during the 2020 Census? If some misclassified LQs could not be converted to the appropriate universe, what were the reasons why?
3. Were there LQs that went through multiple conversions during the 2020 Census? If so, why?
4. What types of data were exchanged between operations to facilitate a conversion during the 2020 Census? Are there additional types of data that are needed to facilitate conversions in the future?

Project: In-Field Enumeration Proxy Procedures

Collecting data from proxy respondents like neighbors or building managers during In-Field Enumeration is essential to get information about hard-to-enumerate households. Proxies also help to verify vacant units and units that do not exist. However, the quality of data collected from proxies may be less accurate than data from other sources and less complete. This project works in conjunction with two research projects, "Case Assignment and Contact Strategy" and "Improving Proxy Responses with Administrative Records," to develop an integrated approach to improving proxy data collection.

Research Topics

- Identify the strengths and weaknesses of the 2020 proxy contact strategy and business rules in the Nonresponse Followup (NRFU) operation.
- Compare the quality and completeness of data collected from proxy respondents to administrative record (AdRec) data.
- Identify ways to improve proxy data quality.
- Develop recommendations for updated proxy procedures, to be tested in the field and refined as needed for the 2030 Census In-Field Enumeration operation.

Research Questions

1. How often were proxy data collected during the 2020 NRFU operation, what was the quality of those data, and how were they used?
2. What changes to procedures could improve the quality of data collected by proxy, reduce respondent burden, and ensure a complete count of people and housing units?

Project: Enhance Methods to Determine Occupancy Status

Decennial census efforts to count the population rely primarily on methods to encourage people to self-respond, followed by in-person interviewing for nonresponding households. To increase efficiency and reduce costs, the 2020 Census leveraged administrative data to enhance respondent-provided information, validate respondent submissions, and to reduce in-person field followup visits. The 2030 Census will continue these 2020 Census innovations while also exploring additional data sources and uses to further streamline census data collection and minimize respondent burden.

Research Topics

- Improve occupancy status resolution with alternate data.
- Utility data and housing-unit occupancy status.
- Real estate data and housing-unit occupancy status.
- Crowdsourced information and housing-unit occupancy status.

Research Questions

1. Can utility data be used to determine whether a housing unit is vacant?
2. Can information displayed on real estate or rental web sites be used to identify vacant housing units?
3. Is it possible to crowdsource reliable data from the public concerning whether a nearby housing unit is vacant?

Project: In-Field Self-Response Options

This project will explore the feasibility of introducing self-response options to in-field enumeration, and thus, provide an alternative(s) to traditional field interviews. In-field self-response options that will be explored include having respondents complete their questionnaire on the enumerator's Census Bureau-issued device and providing respondents with a quick response (QR) code to complete their questionnaire on the respondent's personal device.

Research Topics

- Questionnaire content for the in-field self-response instrument.
- Respondent burden, privacy, and engagement.
- New field procedures and training.
- Technical aspects of the in-field self-response option.
- Feasibility and functionality of utilizing QR codes.
- Risks with providing an in-field self-response option.
- Public interest in an in-field self-response option.

Research Questions

1. How should in-field self-response options affect enumerator field procedures and training?
2. How should in-field self-response enumeration be classified?
3. How should quality assurance be conducted for in-field self-response cases?
4. Which case types should the in-field self-response option be eligible for?
5. Should both in-field self-response options be provided to the respondent, in addition to a traditional interview? What are the factors that may influence which options should be presented to respondents?
6. What are the security measures and device requirements for handing a Census Bureau-issued device to respondents?
7. Should the enumerator check the validity and completeness of response data when the respondent completes their questionnaire on the enumerator's Census Bureau-issued device?
8. How should QR codes be provided to respondents?
9. Are there device or connectivity requirements for the use of in-field self-response options?
10. What ISR-specific security measures are needed to allow the QR code to navigate the respondent directly to the case for their specific address or Census ID?

Project: Household Roster Revision

The decennial census residence criteria say that people should be counted at their usual residence, which is defined as the place where they live and sleep most of the time. If people do not have a usual residence, or cannot determine one, then the residence criteria say they should be counted where they are staying on Census Day. Past research has found a discrepancy between how the Census Bureau uses terms to define who should be included on a roster for an address and how respondents interpret the terms. Respondents have inherent ideas about what "usual residence" means that do not correspond to the Census Bureau's definition or even necessarily a person's physical presence.

The overall project results should demonstrate that the new rostering procedure yields within household coverage that is more consistent across members of the same household and over time. The new procedure will also require fewer, if any, additional instructions but rather guide respondents through a series of simple questions about who is sleeping at the address and any other places those people may be sleeping.

Research Topics

- Household roster question in the 2030 Census.
- Improving within household coverage.
- Correct enumeration of complex, dynamic households.

Research Questions

1. Are respondents able to interpret the new rostering questions as intended during cognitive interviews?
2. Are they able and willing to report who is sleeping at the household?
3. Are they able and willing to provide information about people who sleep somewhere else some of the time?
4. Does the new rostering procedure yield an accurate household roster, as measured by an additional series of undercount and overcount questions during pretesting?
5. During debriefing what do respondents think of the new rostering procedure?
6. Are they able to produce the same roster during a reinterview at a later date?
7. In households with multiple members, is the same roster produced by different household members at higher rates than with the 2020 Census rostering procedure?
8. Is there a lower rate of rostering discrepancy over time between the new procedure and the 2020 Census rostering procedure?
9. Is there a difference in how the new procedure performs for households with historically undercounted characteristics or with people more likely to be duplicated in the census?

Project: Research to Improve Communications, Messaging, and Advertising Efforts

A successful communications and advertising campaign would increase self-response and improve data quality for the 2030 Census. The underlying goal of this set of projects is to conduct research that will inform a data-driven campaign. In this project, four subprojects are proposed that each contribute to this overall goal and will inform both the design and content of the 2030 Census campaign. Each of the subprojects aims to build on the successes of the 2020 Census campaign as well as introduce new methods and strategies to improve upon those past successes. Subprojects aim to inform the 2030 Census communications campaign overall, as well as develop outreach strategies for historically undercounted populations (HUP).

Research Topics

- Measure census barriers, attitudes, and motivators early and often.
- Understand barriers and information consumption habits among HUP.
- Identify and motivate response among HUP.
- Design effective targeted advertisement strategies.

Research Questions

Measure census barriers, attitudes, and motivators early and often

1. How do perceptions of the Census Bureau, decennial census, data privacy and confidentiality, and the federal government vary across demographic groups throughout the decade leading up to the 2030 Census?
2. What communications strategies and data sources can be leveraged to reflect these perceptions, barriers, behaviors, and backgrounds to motivate response in the 2030 Census communications and advertising campaign?

Understand barriers and information consumption habits among HUP

3. How do attitudes toward data privacy, confidentiality, and institutional trust among HUP affect the propensity to self-respond to the census?
4. What is the relationship between information consumption habits and attitudes toward data privacy, confidentiality, and trust among HUP?
5. Are HUP aware of the census results and their benefits during intercensal years?
6. What type of messages, channels, and formats can efficiently help decrease any concerns related to privacy, confidentiality, and institutional trust among HUP?

Identify and motivate response among HUP

7. Which methods and resources can be used to accurately identify households with a low propensity to respond?
8. How can the 2030 Census campaign be designed to motivate these populations to respond?

Design effective targeted advertisement strategies

9. How can target audience groups be adjusted to yield increased respondent cooperation and engagement in the 2030 Census? This includes defining more nuanced demographic target audiences, and reassessing methodologies to create audience segments and census mindsets.
10. Which data sources, channels, and targeted outreach strategies can be built into the 2030 Census campaign that would yield increased respondent cooperation and decreased costs? Strategies may include different outreach platforms, strategic connections with community partners, and leveraging a broad array of data sources.

Project: Content Research for Island Areas Censuses Questionnaires

This project enhances the questionnaire content developed for the 2030 Island Areas Censuses (IAC) questionnaires by ensuring it is relevant and meaningful to people living in the Island Areas (American Samoa, the Commonwealth of the Northern Mariana Islands, Guam, and the U.S. Virgin Islands) and ensuring that it remains consistent with the high-quality content fielded in the 50 states and District of Columbia. This project will thoroughly review findings from previous research and analysis and makes recommendations for new content research through pretesting. It will also thoroughly review lessons learned to identify and prioritize opportunities to refine or improve 2030 IAC questions.

Research Topics

- Island Areas Censuses.
- Stakeholder feedback.
- Content pretesting.

Research Questions

1. Do the questions included in the Island Areas Censuses meet the data needs of the Island Areas? Are there any unmet needs that require new questions? Are there any unmet needs that require modifications to existing questions?
2. Do changes to Island Areas Censuses questions resolve issues documented in lessons learned?
3. Are improvements to ACS questions relevant to Island Areas respondents?
4. Should content changes with potential positive impacts for Island Areas respondents be fielded, even if they are not implemented on the ACS?
5. Should content changes with potential negative impacts for Island Areas respondents be fielded, even if they are implemented on the ACS?
6. Are additional adaptations needed to modify ACS content enhancements for fielding in the Island Areas Censuses?
7. Is it feasible to conduct virtual pretesting with Island Areas respondents?

Project: Targeted Quality Improvement

The purpose of this project is to design and test a Targeted Quality Improvement (TQI) follow-up operation that would replace and expand on the scope of the 2020 Census Coverage Improvement follow-up operation. This will address a wider variety of data quality issues besides coverage (undercount and overcount), including item nonresponse or other data discrepancies, while also minimizing respondent burden by using targeted question pathing and administrative records modeling procedures. The data quality issues of interest include, but are not limited to, item nonresponse, data inconsistencies such as differences between age and relationship (i.e., a two-year old is identified as a parent), and differences between the respondent-provided population count and the number of people on the roster for whom data are provided.

Research Topics

- Coverage improvement.
- Identifying and resolving data quality issues and data inconsistencies during data collection.
- Administrative records modeling and characteristic substitution and edits.

Research Questions

1. Can potential data quality issues be identified in near real time?
2. Can potential data quality issues be addressed with the respondent during the data collection period [during In-Field Enumeration], such that the information can be incorporated into processing in near real time?
3. Can an instrument be built that tailors the set of follow-up questions to the reason the case is sent to the TQI operation?

Project: Improving the Coverage of Young Children in the 2030 Census

This project uses quantitative and qualitative approaches to understand why young children ages 0 to 4 are persistently undercounted in the decennial census. It also explores whether there are operational changes that can be implemented to try to limit undercount. The Census Bureau made improving the count of young children a priority for the 2020 Census. However, young children were still undercounted at a higher rate than other age groups in the 2020 Census, and at a slightly higher rate than in 2010. There has been extensive research by the U.S. Census Bureau and others on this topic. While we don't know exactly why young children are missed at higher rates than other age groups, we do know a lot about the various factors associated with the undercount of young children. In 2022, the Census Bureau formed the Young Children Working Group, which is a cross-directorate team focused on the undercount of young children and improving data quality on this population. The Working Group has developed 11 subprojects to help us understand why young children are undercounted and how we can improve the coverage of this population in the 2030 Census. The Working Group is taking a multiprong approach with the methods ranging from machine learning and administrative records linkage to cognitive interviews and community-based participatory research.

Research Topics

- Undercount of young children.
- The characteristics of young children compared to other age groups.
- Different types of coverage error and their impact on the count for young children.
- Using community-based participatory research, focus groups, and other methods to understand why young children are undercounted in the census.

Research Questions

Characteristics of Young Children

1. What are the characteristics of young children counted in the 2020 Census and ACS?
2. How is household structure related to coverage errors in the 2020 Census? How do complex living situations vary by age, race, and Hispanic origin?

Census Operations

3. What proportion of young children had their age imputed in the 2020 Census and how does this compare with the 2010 Census? How accurate were the age imputations?
4. How can we improve the wording for the age and date of birth questions for the 2030 Census to make them clearer for both English and Spanish speaking respondents?
5. How do the operational metrics for households with young children compare to the operational metrics for other households?
6. How many young children were added through the 2020 Census AR Nonresponse Followup operation? What are the characteristics of households with young children that were enumerated with administrative records?

Supplementing Existing 2030 Design Plan Projects

7. What concerns did people who called Census Questionnaire Assistance have about young children or other coverage issues?
8. What are the barriers, attitudes, and motivators that people have about reporting young children in their Census response?

(continued on next page)

(continued)

Project: Improving the Coverage of Young Children in the 2030 Census

Research on Why Young Children are Missed

9. What are the reasons why young children and other historically undercounted populations may be missed in the Census?
10. What privacy concerns do respondents have about including young child in a government survey and why a respondent may conceal that there are young children living in the household?
11. What are the characteristics of mother who are missed in the American Community Survey (ACS)?

Project: Enhance External Engagement

Enhancing external engagement is critical to the success of the 2030 Census. The Census Bureau relies increasingly on external stakeholders at every geographical level to organize and carry out activities supporting decennial census operations. Identifying and engaging organizations including those that serve historically undercounted populations is an especially important part of these efforts. This project will explore ways the Census Bureau can equip stakeholders to increase their capacity, identify gaps, and leverage networks to further activate resources and improve response. This approach requires place-based testing of concepts such as the "power of convening" -- the process of bringing external stakeholders together to determine structures, systems, capacities, and tools needed to conduct a better enumeration of the U.S. population.

Research Topics

- Information forthcoming.

Research Questions

1. Information forthcoming.

EA 1: Data Collection Post Enumeration Survey (PES)

Project: Modernize PES Data Collection, Introduce Automated Operations, and Explore Internet Self-Response and Alternative Data Sources

The Post-Enumeration Survey (PES) estimates the coverage error of the census. Although the Census Bureau undertakes extensive efforts to accurately count everyone in the decennial census, sometimes people are missed or duplicated. Census errors can result in a smaller or larger population count than the actual number of people. The Census Bureau estimates the population size using the PES. This project focuses on redesigning the PES.

Research Topics

- Reduce dependence on field data collection.
- Automate field data collection tools.
- Automate training for field and matching operations.
- Investigate administrative records and alternative data sources.
- Investigate machine learning algorithms for record linkage.

Research Questions

1. How can the data collection for field operations be timely and allow data to flow into the matching operation in near real-time?
2. What improvements can be made to the data collection solution, systems, and procedures to improve data quality?
3. What is the data collection solution(s) for housing unit and person?
4. What is the most efficient way to collect data such that minimal processing is needed to the data for matching operations?
5. How can the questionnaires be redesigned to collect data more efficiently?
6. How can it be ensured that PES interviewers stay in the correct geographic boundaries (i.e., map display for field data collection, improved training, or some other strategy)?
7. Is online training an effective method to deliver data collection and matching content? Should this be included in the development of the automated solution? Can consistent training methods be developed across all operations?
8. To what extent can fieldwork for a post-enumeration survey be replaced with other data sources (such as administrative records or web-scrapped data)?

Project: Create Computer-Assisted Post-Enumeration Survey Clerical Matching Software

The Post-Enumeration Survey (PES) estimates the coverage error of the census. Although the Census Bureau undertakes extensive efforts to accurately count everyone in the decennial census, sometimes people are missed or duplicated. Census errors can result in a smaller or larger population count than the actual number of people. The Census Bureau estimates the population size using the PES. This project explores various avenues to reduce the clerical matching workload, improve the person and housing unit computer matching software, and make the clerical matching activities easier and more efficient.

Research Topics

- Use guided prompts in clerical matching.
- Incorporate machine learning in matching tasks.
- Improve computer matching software.
- Establish a training and testing database.

Research Questions

1. To what extent can the clerical matching software system be improved by including prompts with questions for the clerical matchers to answer to guide the PES coding?
2. To what extent can a list of questions and skip patterns be created to guide clerical matchers to probable codes to assign, e.g., match codes and status codes, for scenarios in clerical matching?
3. Would including prompts with questions for the clerical matchers to answer result in as accurate or more accurate results than 2020 PES clerical matching results? Would clerical matchers work more efficiently with the prompts?
4. To what extent can machine learning algorithms be used to reduce clerical matching (person and housing unit), residence status coding, and geocoding workloads?
5. To what extent can alternative record linkage algorithms be used to replace the current computer matching software system?
6. To what extent can high-quality alternative or additional data variables and maps be used in machine learning to get more high-quality links to reduce the clerical matching workload?
7. To what extent can machine learning be used to reproduce the match codes assigned by the clerical matchers?
8. How can machine learning or other matching algorithms be used to alert the clerical matchers to possible errors, real-time during the clerical matching, or to suggest appropriate codes?
9. How can current geography indicators be improved, augmented, or replaced with more accurate indicators of geographic distance, e.g., distance in miles between addresses, to create a measure of confidence in the links made between addresses?
10. How can nationwide matching be refined and common names made harder to link to reduce false matches?
11. How can the list of invalid names, like "Female Number 1" and "No One Home," be expanded and used to improve PES person computer matching results by reducing matches with false names? How can an automated system identify names that are complete but insufficient for follow-up?
12. How can data standardization for computer matching be refined to be more like census standardizations to improve PES person computer matching results?

(continued on next page)

(continued)

Project: Create Computer-Assisted Post-Enumeration Survey Clerical Matching Software

13. How can a training and testing database with realistic data that reflects the complexities of real-life geography, housing unit and person demographics be created and consistently updated?
14. What are the most important concepts that are required for training and testing to ensure the success of PES?

EA 2: Modernize Group Quarters Enumeration

Project: Stakeholder Engagement, Living Quarters Definitions, and Service-Based Enumeration

This project includes three sub-projects. The first involves outreach with multiple stakeholders to enhance all aspects of group quarters (GQ) operations. Key activities include engaging federal, state, and other organizations, working with umbrella organizations for group quarters and transitory locations data sources, establishing agreements with external stakeholders, engaging organizations to establish criteria for survey participation, developing a one-stop group quarters and transitory locations webpage, along with a service-based enumeration website, and group quarters and transitory locations-specific advertising & awareness.

The second sub-project involves working with internal and external stakeholders to update the decennial census (and American Community Survey) living quarters definitions as well as definitions for each of the specific group quarters and transitory location types.

The third sub-project is focused on improving the Service-Based Enumeration (SBE) operation by teaming with the Housing and Urban Development and the Continuum of Care Centers to learn data collection techniques for collecting data on people experiencing homelessness. This project also researches new living situations, such as vehicle residency, that people experiencing homelessness is using and the best ways to locate and count them at these places.

Research Topics

- Improve opportunities for gaining access to group quarters and transitory locations.
- Establish Memoranda of Agreement or Memoranda of Understanding.
- Improve sources for all GQE, SBE, and transitory locations data.
- Gain better understanding about state and federal laws.
- Develop a nationwide and targeted advertising plan.
- Develop webpages for stakeholders.
- Explore other self-response alternatives for GQ residents.
- Extent to which updating and revising living quarters definitions provides clarity for all users.
- Develop living quarters definitions that reflect changing living situations.
- Develop, improve, and modernize the Service-Based Enumeration operation.
- Add additional service-based locations to the workload.

Research Questions

1. How can outreach to external stakeholders support the improvement of GQ and TL frame building and data collection for the 2030 Census?
2. What partnership activities can we initiate to engage external stakeholders throughout census tests?
3. What are the most effective methods to create public awareness to promote and encourage participation for the enumeration of people residing in all GQs and TLs?
4. Can additional external stakeholders be identified by FY 2024?
5. Which living quarters definitions can be improved or clarified? How? Should any group quarters or transitory location type codes be added, removed, or merged?
6. How do the housing unit and transitory location definitions need to be revised to reduce confusion or overlap with certain types of group quarters?

(continued on next page)

(continued)

Project: Stakeholder Engagement, Living Quarters Definitions, and Service-Based Enumeration

7. Will the proposed changes to the living quarters definitions lead to significant work to align the MAF with revised definitions?
8. Will any of the proposed changes to the living quarters definitions create more problems than they solve?
9. How do anticipated providers of group quarters and transitory locations data classify these types of structures? Do their definitions align with ours?
10. Did the website designed for local community leaders and service providers to submit targeted non-sheltered outdoor locations and other SBE group quarters types to the Census Bureau for enumeration improve frame building and data collection?
11. Did deploying a mobile application and implementing modernized procedures improve data collection for the SBE operation?
12. What methodologies used for the FY 2024 small-scale test can be used to inform the Census Test in FY 2025 through FY 2026?

EA 2: Modernize Group Quarters Enumeration

Project: Group Quarters and Transitory Locations Frame Sources

This project has three main components that could help improve the Group Quarters (GQ) Frame: Business Register and Administrative Records, INFO-COMMS, and Web Scraping. These projects support the "Not a Housing Unit Working Group" under the Census Bureau's Geographic Support Program by researching data sources from the Business Register, Administrative Records, Administrative Response Records, Facility Records, etc., and reviewing 2020 INFO-COMMs for information about group quarters. This project will also research if Web Scraping and Associative Information Extraction for GQs can help improve the GQ Frame. For Enumeration at Transitory Locations (ETL), this project explores ways to create an accurate, reliable, and updated transitory location (TL) frame prior to the decennial enumeration benchmark, exploring sources for use to regularly update the frame and include continuous improvements to the TL frame data throughout the decade leading up until the 2030 Decennial Census.

Research Topics

- Research administrative records that can be used to improve frames.
- Exploring additional datasets such National Student Clearinghouse.
- Find new sources to update frames continually throughout the decade.
- Evaluate all identified datasets for fitness for use. -If these additional datasets are deemed beneficial for use in decennial operations, then work to coordinate with the Geographic Support Program (GSP) to incorporate these datasets into decennial IT systems

Research Questions

General

1. Can we identify sources that will improve the GQ and ETL frames? This overarching research question will be answered by addressing the following research questions for each of the main components of the project

Business Register and Administrative Records Research:

2. How well do particular North American Industry Classification System (NAICS) codes identify companies which manage group quarters and transitory locations?
3. For NAICS codes that have many, but not all, companies which manage GQs and TLs, is there a way to identify or predict with a model which ones manage GQs?
4. How complete is the data on the Business Register and could it be used as a reliable source to update the GQ and TL universe?

INFO-COMMS Research:

5. How did FLD staff address access and cooperation issues with GQ facility contacts or Point-of-Contacts?
6. How commonly did addresses change, GQ, HU, or TL conversions, and refusals occur?
7. Can the information obtained from the INFO-COMM be used to update the Master Address File (MAF)?
8. Were changes on INFO-COMMS reflected in address outcomes in post-processing?

Web Scraping and Associative Information Extraction for GQs and TLs:

9. Which type of GQs are best for web scraping and data extraction? Meaning do their websites contain information such as unit or occupancy information useful for the GQ frame and do the sites permit web crawling?

(continued on next page)

(continued)

Project: Group Quarters and Transitory Locations Frame Sources

10. How complex does the program code need to be in order to identify the pages for the desired data?
11. Are there website formats that will require adjusting the web scraping program code for web scraping and data extraction?
12. What are the differences between college-owned student housing facilities versus privately owned student housing locations?
13. How accurate is the status of GQs and TLs on Google in regards to 'operational' vs. 'closed' status?
14. Can web scraping be used as a source to update the TL frame? What websites or resources could be used to update the TL frame via web scraping? What information is available on the websites and do the websites offer a complete or accurate picture of various TL types where web scrapping is a viable resource for frame updates?

Project: Improving Group Quarters Validation

This project will reinstitute and enhance the 2010 Group Quarters Validation (GQV) operation to verify and improve the data in the group quarters (GQ) and transitory location (TL) universe. The GQV operation received the workload of addresses that were identified as Other Living Quarters (OLQs) from the Address Canvassing operation. OLQs addresses were addresses that were not commonly recognized as a Housing Unit. OLQs include such places as group quarters (e.g., skilled nursing facilities, group homes intended for adults, and college and university student housing) and transitory locations (e.g., campgrounds, hotels, and motels). The 2010 GQV operation significantly reduced and eliminated duplicates in 2010. GQV identified 42.68 percent (873,764) of the OLQs as non-valid living quarters which yielded 874,081 total units. Non-valid living quarters were those addresses that GQV found to be either deletes, duplicates, or non-residential. To support this address validation project, hiring and training personnel to be GQ and TL specialists is required at HQ and in the field.

Research Topics

- Validate a more accurate universe for the GQ and TL operations.
- Determine how the GQV operation would correctly classify all OLQs.
- Determine whether GQV should be entirely an in-field operation.
- Identify deletes and duplicate addresses.
- Accurately link GQs to the parent organization or facility.
- Determine the QC criteria when adding GQs or TLs.
- Create training for the staff.
- Enhance listing procedures.

Research Questions

1. Does reinstituting the GQV improve the identification, classification, and geocoding of Living Structures that include GQs, TLs, and HUs?
2. Does reinstituting the GQV facilitate the linking of GQs to parent organizations or facilities?
3. Does the GQV operation verify an inventory of GQs, TLs, and housing units (HUs) in the correct location?
4. Does the GQV operation geocode the GQs, TLs, and HUs addresses correctly?
5. Does the GQV operation accurately classify the structure at each address?
6. Does the GQV operation successfully link the GQs to the parent organization/facility?

EA 2: Modernize Group Quarters Enumeration

Project: Automating Group Quarters and Transitory Locations Operations

This project will focus on the automation of data collection solutions for use during in-field enumeration and reinterview for the broad range of Group Quarters (GQ) operations, including Military, and Enumeration at Transitory Locations (ETL) operations. Automated data collection will reduce or eliminate the use of paper questionnaires used in the field. Automation also reduces the amount of paper forms printed, shipped, stored, and processed by the Census Bureau during the Decennial Census. Through automation, field staff can manage case assignments and data collection in real time. Automation can enable real time respondent data analysis for quality assurance and progress tracking. The project will accomplish work across multiple fiscal years, starting in FY 2022 and ending in FY 2030.

Research Topics

- Research automated data collection solutions for the following operations: Group Quarter/Transitory Location Validation (GQV), Group Quarters Advance Contact (GQAC), Group Quarters Enumeration (GQE), Service-Based Enumeration (SBE), Transitory Locations Advance Contact (TLAC), Enumeration at Transitory Locations (ETL).
- Research the best type of automation.
- Use existing technology where possible.
- Research best options for online training.

Research Questions

1. Are electronic enterprise solutions an effective data collection method for GQ and TL operations?
2. What is the best approach to creating an automated Data Collection or Reinterview instrument?
3. Is online training an effective method to deliver learning content? Should this be included in the development of the automated solution? Can consistent training methods be developed across the enterprise for all operations?
4. What will be the minimum/optimal quarters and transitory locations data collection standards and criteria?

EA 2: Modernize Group Quarters Enumeration

Project: Create Internet Self-Response Option for Non-Institutional Group Quarters

This project is to develop an internet self-response (ISR) instrument specifically for certain types of non-institutional group quarters (GQ), such as college student housing or military barracks, to support the 2030 Census Group Quarters Enumeration (GQE) operation. The GQE-ISR instrument will be an online tool that allows the respondents living in certain non-institutional GQs to complete and submit their Individual Census Questionnaire (ICQ) electronically. The GQE-ISR will securely capture and transmit respondent data and will use unique identifiers to ensure the respondent is counted at the appropriate GQ. The GQE-ISR will leverage the design of the 2020 Census ISR instrument for housing units, but it will be developed as a separate instrument that can work in tandem with the housing unit ISR instrument for the 2030 Census.

Research Topics

- GQE-ISR instrument designed for specific GQ types.
- Ability to link respondents to the group quarters.
- Real time communication with other solutions.
- Effective hyperlinks.
- Determine if people know how to respond.

Research Questions

1. What happens if GQE-ISR is not completed for residents at specified GQ types? How does a case get remapped to other types of collection?
2. How do we track if all the residents of the GQ are enumerated if the GQ uses GQE-ISR in order to know when that GQ's enumeration is complete?
3. What are optimal communication modes and messages to ensure GQE-ISR participation by GQ residents?
4. Can GQ administrators be leveraged to provide information about GQE-ISR to residents and motivate their residents to complete GQE-ISR? If so, what kinds of information, formats, etc. would they want us to provide?
5. Can only GQ resident respondents use the GQE-ISR? Or is it also a tool for administrators at the in-scope GQs?
6. How are nonresponding GQ residents handled? Do they get additional contacts? Does a GQ administrator respond using GQE-ISR as a proxy?
7. How does ACS and other surveys of GQ residents handle respondent contact and nonresponse follow-up activities?
8. What if there are more or fewer GQE-ISR GQ resident responses than the administrator said was at the facility? Is there a real-time processing check that could initiate a contact to follow up on discrepancies?

EA 2: Modernize Group Quarters Enumeration

Project: Electronic Response Enhancements to Facilitate Group Quarters Administrators' Response

This project includes a number of key enhancement areas to modernize data collection by improving group quarters (GQ) administrators' ability to respond electronically with accurate, complete data in multiple formats. This project will leverage existing enterprise solutions for web-based data collection. It supports the following 2030 Strategic goal 1 of a complete and accurate census and objective 1.1, which states that all living quarters are associated with an address.

Research Topics

- Identifying multiple GQs combined under one identification.
- Efficient method of uploading different templates.
- Real time generation of GQ and client identifiers.
- Real time generation and reset of personal identification numbers (PINs).
- Real time acceptance of vacant GQs.
- Real time integration with field operating systems.

Research Questions

1. How we can develop an electronic response data collection solution to accept raw GQ administrative records and convert them into acceptable and usable formats?
2. Can a web-based GQAC operation provide opportunities to receive GQAC data directly from GQ or TL administrators to provide efficiency and accuracy in the collection of data at the GQ level?
3. Will an optimally-designed web-based GQAC operation improve the quality and completeness of GQ-level records?
4. Is the collection instrument or system able to identify and flag cases where multiple GQs are submitted under one GQ identification number, and then allow the GQ administrator to correct the error in order to enumerate each separate GQ accurately?

EA 2: Modernize Group Quarters Enumeration

Project: Group Quarters Count Expectation Modeling to Ensure Data Quality

Errors in individual group quarters (GQ) counts, especially for larger GQs, are highly visible to the public and could adversely impact the perceived quality of the census. Large GQs are often the only living quarters in their tabulation census block. Consequently, information suggesting that a GQ has too low or high population, or zero, would be evident in the PL 94-171 Redistricting Data Summary File.

Research Topics

- Investigate procedures that provide reasonable counts for each GQ.
- Simulate GQ information from the 2020 Census.
- Evaluate methodologies to model the GQ counts.

Research Questions

1. Is it possible to supply managers with expected GQ counts prior to census operations so they can be informed and, if appropriate, take action during the GQ data collection?
2. If no count is obtained in the GQ Enumeration operation, is the modeled count accurate enough be imputed for the GQ?
3. How do we indicate the potential variability in the count?

EA 3: Integrate Data Processing with Data Collection in Real Time

Project: Centralized Decennial Address Control (CDAC)

This project will propose a central, readily accessible decennial address repository that allows for near real-time updates and is the final authority on decennial address matching, addition, and update of addresses, and assigning permanent decennial Reporting Unit IDs. A central decennial address repository updated in near real-time would allow for flow-based response processing, as opposed to the 2020 Census workflow in which all updates had to be received, processed into the Master Address File/Topologically Integrated Geographic Encoding and Referencing System (MAF/TIGER), benchmarked and output in products (e.g., Final Collection MAF Extract - MAFX) before response processing could begin. That workflow, used in the last three censuses, forces most of the work towards the end of the decennial year and limits the amount of analysis and remediation that can be done in time to meet the apportionment deadline. In addition, a currently updated decennial address repository can improve matching results by always matching to a current universe, as opposed to the 2020 Census, where Real-Time NONID Processing matched against an address benchmark copy that was not current.

Research Topics

- Near Real-Time Matching.
- Generating address update transactions from Collection Events.
- Address standardizer improvements.
- Address visualization tools.
- Streamlining correction or repair of addresses and geocodes.
- Flow-based address processing.

Research Questions

1. Can near real-time matching be conducted against a live address repository (as opposed to using a benchmark and building associated auxiliary matching tables, as was done for the 2020 Census)?
2. How will address transactions be implemented based on a flow of collection events sent to CDAC?
3. Will potential changes to Internet Self-Response address collection screens require new address standardization tools or methods?
4. How can this proposal support improvements in the analysis and remediation of address issues as they occur, including housing units, transitory locations, and group quarters nonresidential conversions.
5. To what extent can time-consuming address corrections or geocoding fixes that could hinder near real-time processing be avoided or streamlined?
6. Can a centralized address repository support flow-based address processing obviate the need for a Final Collection MAFX and NonID Feedback Table (NIFT)?

EA 3: Integrate Data Processing with Data Collection in Real Time

Project: Near Real-Time Race and Ethnicity Coding

Coding of the race and ethnicity question during the 2020 Census was conducted in batches or "waves" which lead to bottlenecks in processing. As part of an overarching near real-time processing platform, the current race and ethnicity coding process will be modernized and pivot away from this batch processing conducted in previous Censuses. Research will investigate what methods can be implemented to improve race and ethnicity coding and how it will be incorporated into the overall near real-time response processing flow.

Research Topics

- Updating the autocoding dictionary in English and non-English languages in a transactional manner.
- Strategy for coding race and ethnicity entries in near real time.
- Strategy for updating the autocoder in near real time.
- Machine Learning as an option for augmenting the autocoding dictionary and perhaps minimizing the case load for National Processing Center clerical coding.
- Developing appropriate analytics for clerical and automated processes.

Research Questions

1. Which race and ethnicity coding functions from the 2020 Census can be done in near real-time?
2. What is the best approach to implementing these functions?
3. How do these processing functions interact with workload management?
4. What proportion of 2020 Census non-English responses sent to response translation were already included in the autocoder?
5. How can these processing functions be used to support real-time data quality monitoring?
6. How can "natural language processing" enhance the coding of race and ethnicity write in entries? Can a "machine learning algorithm" be integrated with the Hispanic Origin and Race Masterfile updates.
7. What process should be implemented to ensure the quality of the residual coding procedure and results?

EA 3: Integrate Data Processing with Data Collection in Real Time

Project: Integrated Performance Analytics

During the 2020 Census, particularly during the data collection efforts, most of the metrics generated were used to inform managers about how an operation was progressing. This project studies if and how the 2030 Census program can create an integrated performance measurement culture. The key deliverables of the project include an integrated performance analytics framework describing the best approaches to achieve a true integrated culture and sample dashboards that can address most of the user measurement needs for the 2030 Census.

Research Topics

- Establish a framework to address the needs of performance management and data analytics.
- Determine a baseline set of requirements and inputs that cover the metrics needed for the 2030 Census (operational and data quality).
- Create a proof of concept that enables management to choose the most efficient and effective design.

Research Questions

1. What is the best performance framework for data analysis and performance management across the 2030 Census program?
2. Could a small number of well-developed and well-tested reports address most key stakeholders?
3. What pre-determined data analytics, metrics, and key performance indicators would be needed for those small number of reports?
4. Can metrics at a granular level of detail be produced without performance impacts?
5. Can the timeliness of the report data be improved to address the varying needs of stakeholders without overburdening system performance?
6. Can a well-defined and well-integrated ad-hoc reporting process be developed to address the reporting needs that arise during production?
7. What are the metrics to allow managers to appropriately assess data quality, operational progress, response sufficiency, outlier analysis, oversight or external expectations, data processing benchmarks towards dissemination deliverables?
8. What dashboards, analytics, etc. are required to track quality and completeness?
9. What metrics and reporting tools are best to inform data quality in real time for collected and administrative records data?
10. Which metrics are well-suited for automated decision-making, and which require human intervention?

EA 3: Integrate Data Processing with Data Collection in Real Time

Project: Decision Point Framework

As part of a near real-time processing system, a framework of various decision points surrounding the census is needed. "Decision points" are necessary to fully document where analytics or other methods could help make a decision, either automated or human, during integrated data collection and processing operations. Documenting decision points early, during the research phase, will allow the program to implement "levels" that can be used to adjust course during testing and in production. Ultimately, this project will build a decision point framework so that future decisions can be categorized and documented, which will guide both the research and eventual test and production operational work in the new real-time data processing paradigm.

Research Topics

- Investigate where analytics or other methods could help us make a decision – either automated or human – during our integrated data collection and processing operations.
- Use recommendations from the parallel projects and bring them together into a framework that can then be implemented.
- Develop a tool to determine how to prioritize these findings into the final production systems.
- Create a comprehensive list of decisions for the program.
- Categorize the decisions into types based on source (person or computer) and other criteria.
- Use the framework to aid in building out the solution systems and ensuring the proper levers are available to the appropriate stakeholders.

Research Questions

1. What set of metrics should be developed for determining when responses are at an appropriate level of quality to be considered final?
2. How do we carry forward our decisions we made in 2020 to help us plan better for 2030?
3. What's the best way to communicate to system developers the information we will need to see in order to make decisions?
4. What kinds of decisions will need metrics that are currently unavailable? How will these be treated?
5. How do we ensure we have a level to pull to make in impact on the program based on a decision?
6. What dashboards, analytics, etc. are required to track quality and completeness?
7. What metrics and reporting tools are best to inform decisions around data quality in real time for collected and administrative records data?

EA 3: Integrate Data Processing with Data Collection in Real Time

Project: Real-Time Analysis of Data

The purpose of this project is to produce metrics during data collection that allow decennial to monitor data quality issues, identify areas of concern, and support operations and management to provide additional metrics and analytics as needed. The Real Time Analysis of Data (RTAD) team will collaborate with other analytics and data quality teams to make sure the results are communicated and acted on when necessary.

Research Topics

- Evaluate the effectiveness of the 2020 Census RTAD results.
- Build dashboards based on operational needs.
- Develop metrics based on regularly produced summaries, maps, or tables available in a dashboard.
- Develop metrics based on automatic operational events that occur when predetermined or statistical thresholds are crossed.
- Develop metrics based on a set of geographies or cases that are output for review when a predetermined or statistical threshold is crossed.
- Develop metrics based on ad-hoc summaries to respond to inquiries.

Research Questions

1. How did the 2020 Census RTAD metrics compare to the same metrics when using post-collection processing data?
2. What were some self-response patterns that occurred in the 2020 Census? What were some NRFU response patterns that occurred in the 2020 Census?
3. What issues did RTAD uncover in the 2020 Census?
4. Were there issues that RTAD did not uncover that were in scope?
5. Can the successes of 2020 Census RTAD be improved upon for displaying 2030 Census metrics in real time?
6. Can RTAD results be integrated to automatically make operational changes?
7. Can preliminary operational data quality metrics be produced before the end of data collection?

EA 3: Integrate Data Processing with Data Collection in Real Time

Project: Near Real-Time Processing

Past censuses have generally performed much of the post-collection processing of response data in a linear fashion after data collection operations ended and the census address universe was finalized. Generally, each census file or product is given a few weeks to a few months to process, as well as time for review and to perform quality control. This is not ideal if issues are found with the data or the processing because it does not allow ample time to address the issues.

One aspect of this project is to build a prototype system that allows many of the processing steps to be done nearly as soon as the data are collected, or for some processing steps, at pre-determined frequencies during the data collection period. By shifting as many post-collection processing activities as possible to near real time, the need for a large-scale post-collection processing operation that takes several months can be minimized. This will also facilitate near real-time monitoring of the data to identify data issues or response processing issues that can be addressed during operations rather than with patches or reruns on the back end. Additionally, identification of data-defined person records and calculation of household size, Usual Home Elsewhere (UHE) processing and nationwide person matching resolution are included in near real time processing research.

Research Topics

- Determine which post-collection processing functions from the 2020 Census can be done in near real time.
- Determine the best approach to implementing these functions.
- Determine how these processing functions interact with workload management.
- Determine how processing functions can be used to support near real time data quality monitoring.

Research Questions

1. When evaluating response data, what should constitute a data-defined person (DDP) record?
2. Research the algorithm used to determine household size on a response. When there is a count discrepancy between respondent provided population count and the number of persons provided on the roster, can alternative sources such as administrative records help resolve the difference or suggest different rules?
3. What are alternative methods for implementing a Primary Selection Algorithm?
4. How can 2020 Census methods to resolve duplication (person-level and address-level) be used jointly and improved upon?
5. What impact does UHE processing have on the final census universe?

EA 3: Integrate Data Processing with Data Collection in Real Time

Project: Quality Improvement for Listing and Enumeration

The goal of this project is to recommend effective and cost-efficient quality assurance techniques for the 2030 Census. This will primarily affect In-Field Enumeration and may also be applied to various Post-Enumeration Survey (PES) operations.

Research Topics

- Document the quality goals for 2030 Census operations.
- Identify strengths and weaknesses in 2020 Census quality procedures.
- Investigate techniques to identify poor-quality work.
- Investigate further uses of administrative records.
- Integrate with 2030 Census field operations.
- Coordinate with other relevant programs.

Research Questions

1. How can potential falsification or poor-quality work be identified and addressed when an interview for a selected reinterview case cannot be completed?
2. Are there options, aside from reinterview, that could be used to assess the quality of the enumerator's work?
3. How can cases be best selected that are potential falsified or are of poor quality for further review? In addition to the administrative records and analytic sampling used in the past, could third-party data, machine learning, data from other census operations, or other techniques to select cases also be used?
4. In what ways can computer matching between production and reinterview data be improved?
5. How can the clerical work required be severely reduce or even eliminated? Can machine learning be used to reduce the clerical workload?
6. If there are new field assignment types (such as sending listers to find only specific addresses), what QC is appropriate for those assignment types?
7. For any listing design, what can be improved for QC listing procedures and sample design to meet the desired quality goals while containing costs?

EA 3: Integrate Data Processing with Data Collection in Real Time

Project: Characteristic Imputation Research

Although the Census Bureau strives to obtain all demographic and housing data from every individual in the census, missing data are part of every census process. To account for this, the 2020 Census enhanced long-established procedures in previous censuses and surveys to fill in these missing characteristics. This project will study improvements and expansion of imputation methods, including using modeling instead of direct assignment, substitution, and allocation as well as developing models for group quarters.

Research Topics

- Explore additional statistical modeling applications to become the main method of imputation.
- Explore statistical applications to enhance quality in the characteristic imputation procedures and further enhance the methods that the Census Bureau uses to account for missing characteristics.
- Evaluate alternatives for implementing these methods as part of the real-time processing initiative.

Research Questions:

1. Can census characteristic imputation make better use of modeling approaches and administrative data to compensate for missing data?
2. The 2020 Census used past census and administrative records to do a direct assignment of characteristics. Instead of direct assignment, can a modeling approach be done instead?
3. How do the results of these imputation approaches compare to the traditional imputation approaches used in the 2020 Census?

EA 3: Integrate Data Processing with Data Collection in Real Time

Project: Improving Proxy Responses with Administrative Records

The overarching goal of this project is to provide recommendations to improve the quality of proxy responses to the census. We seek to build on the 2020 Census where we used high-quality administrative record (AR) responses in lieu of a proxy interview. The improvement is attained by increased use of AR data, either while the proxy interviews are being completed during the In-Field Enumeration operation or in the post-processing operation.

Research Topics

- Analyzing scenarios for which AR data might be an improvement over proxy responses from the field operation.
- Providing use cases indicating how AR would improve the quality of proxy responses.

Research Questions

1. At the address level, how complete were proxy responses for the 2020 Census as compared to what was known in AR?
2. In what situations can we inform field staff that a proxy response is less preferred than the closeout-quality AR data?
3. Given the roster supplied by the proxy remains unchanged, in what situations should a proxy enumeration by supplemented by AR?
4. In what situations should a proxy enumeration be supplemented by AR in terms of roster creation? What other sources can we incorporate that were not available to analyze this question for the 2020 Census?

EA 3: Integrate Data Processing with Data Collection in Real Time

Project: Response Sufficiency

This project will use response data, paradata, and any relevant assessment reports and operational lessons learned to evaluate how well 2020 Census response sufficiency criteria worked, in relation to both ensuring high-quality census data and minimizing respondent burden. The project team will research how the response sufficiency criteria can be improved for the 2030 Census while working within the vision of near real-time post-collection processing and data quality issue resolution.

Research Topics

- Designing response sufficiency criteria for all data collection operations and modes.
- Exploring whether response sufficiency criteria should vary by mode, operation, source, stage of data collection, availability of administrative records, or other factors.
- Deciding what data items each data collection instrument should require a respondent to provide through the implementation of hard edits.
- Designing business rules for what specific actions will occur in various scenarios after the sufficiency criteria are applied to a response.

Research Questions

1. What lessons are learned about hard edit rules and response sufficiency criteria by analyzing 2020 Census response data and paradata?
2. What can be learned by examining response break-off patterns?
3. How burdensome to the respondent or user are hard edits for particular data items in each of the data collection instruments?
4. What data items are important for ensuring the accuracy and quality of Census data products?
5. What are the key data items needed for matching and other processing steps?
6. Do certain data collection modes or operations acquire more accurate or complete response data than others?
7. What types of sources provide the most accurate or complete response data?
8. Can the existence of quality administrative records data for particular persons in a household response be verified in near real time, right after a response is received?

EA 3: Integrate Data Processing with Data Collection in Real Time

Project: Near Real-Time Count Review Operation

During the 2020 Census, the Census Count and File Review (CCFR) component of the Count Review Operation evaluated responses and subsequent data files for demographic reasonableness and to verify that post-processing logic had been properly applied. This independent analysis resulted in the identification of hundreds of potential anomalies in the 2020 Census data review files. Several of these anomalies resulted in programmatic code corrections and file reruns, data re-collection and new imputation and unduplication processes for group quarters (GQs), and patches to correct erroneous collection and processing results. The motivation for the Near Real-Time Count Review Operation project is based on the Census Bureau's experience identifying data anomalies during the 2020 Census. This project will study conducting near real-time data review to alert Decennial operations of potential issues that can be addressed during data collection.

Research Topics

- Incorporate near real-time data review into the subject matter expert demographic reasonableness and processing data review for the 2030 Census.
- Updating demographic reasonableness, processing, and count review methods to evaluate data quality and reasonableness in near real time.
- Using the Near Real-Time Count Review Operation to alert census data collection operations of potential data anomalies that should be addressed while data collection is ongoing.
- Identify data review steps that need to be conducted after data collection has ended and data processing is complete.

Research Questions

1. How can the CCFR component of the Count Review Operation be improved to allow for data review to be conducted in near real time?
2. How can the Census Bureau conduct data review while ongoing updates to response data (from housing units, GQs, and persons) and addresses are occurring in near real time?
3. What is the appropriate review process for identifying anomalies and determining demographic reasonableness of population counts and characteristics that require an accumulation of data?
4. How can the 2030 Census projects for Integrated Analytics and Real-Time Analysis of Data be used to identify potential anomalies?
5. Does a separate review tool, like CRAVA for the 2020 Census, need to be developed to assist Census Bureau analysts during data review?
6. How can Census Bureau analysts access the data in near real-time to examine the underlying microdata and run ad-hoc queries?
7. Which review steps need to be conducted after data collection has ended and final response processing is complete?
8. How should near real-time data review interact with workload management to address potential anomalies while data collection is ongoing?

EA 4: Streamline Operational Support Infrastructure

Project: Reimagining Regional Census Centers and Area Census Offices

The Decennial Census Management Division's, Reimagining Regional Census Centers (RCCs) and Area Census Offices (ACOs) project will support the Decennial Logistics Management operation (DLM). DLM coordinates space acquisition and lease management activities for the Regional Census Centers (RCCs), Puerto Rico Area Office (PRAO), Area Census Offices (ACOs), and the National Processing Center's (NPC) additional logistics warehouse, specifically rented for kit assembly and shipping activities.
The Reimagining Regional Census Centers (RCCs) and Area Census Offices (ACOs) project will determine the concept, design, and functionality for the 2030 RCCs and ACOs (i.e., number and type of RCCs and ACOs, full service ACOs vs ACO depots, virtual ACOs, etc.). The preliminary optimal design or set of designs will support the 2030 Census enhancement area, Streamline the Operational Support Infrastructure.

Research Topics

- Centralizing payroll functions to a few ACOs versus having them in every ACO as we did in 2020.
- Creating ACO Depots where a small staff are available to manage and distribute supplies, IT equipment, promotional materials, and other essential items.
- Establishing virtual ACOs versus brick-and-mortar ACOs in areas where a management and supervisory need exists to manage the work, but a physical presence is not necessary.
- Reducing the footprint of the physical RCCs and ACOs to accommodate a smaller number of staff physically present at the offices each day.

Research Questions

1. How can we reimagine what the regional census centers and area census offices will look like for the 2030 Census (i.e., full service ACOs versus ACO depots, virtual ACOs, etc.)?
2. What preliminary conceptual designs and functionality can be identified for RCCs and ACOs by the end of FY 2024 (i.e., full service ACOs handling all administrative functions versus smaller ACOs and ACO depots, RCCs and ACOs designed with teleworking in mind, etc.)?

Project: Human Resources and Onboarding Policy Decisions

This project will be a collection of key stakeholders and decision makers from across the bureau who are integral to making important decisions regarding how we will recruit, select, conduct background checks, and train Schedule A employees for the 2030 Decennial Census. This group will be responsible for identifying and providing guidance on critical human resources (HR) and Onboarding policy.

Research Topics

- Recruiting, selection, and hiring.
- Background Checks.
- Fingerprinting and Badging.
- Training.
- Pay and benefits.
- Time and Expense.
- Safety and Security
- Employment and Hiring Waivers.

Research Questions

1. Can we include the Area Census Office (ACO) management staff with the field staff recruiting, hiring and pay solutions within the ACO?
2. Can a waiver be obtained to include travel expenses in a DAPPS-like system for 2030?
3. Can we figure out a way to pay Remote Alaska and UE field staff in advance for their travel to get fingerprinted and trained?
4. Which system will Decennial field staff use to enter Time and Expense for operating or personnel and payroll?
5. How can we shorten the Application and Application Assessment or eliminate all together?
6. Can NPC leverage a DAPPS-like system for hiring their short-term decennial staff at NPC's Decennial Logistics warehouse?
7. Do we plan for a paper backup for 2030 (all systems) (not Remote Alaska)?
8. Do we need an English Proficiency Test? Can we update it? How?
9. Do we have to ask the selectee whether they have a license to operate a vehicle?
10. What will be the background check time frame? Do we need 60 days?
11. Will we need waivers for Foreign Nationals (Foreign Language Translator) and what is the process for obtaining those waivers?
12. Will we utilize and incentive program for field staff (e.g., travel awards, production award, retention award), and what is the process to obtain approval and implement the program?
13. How much telework in ACOs? (i.e. CFMs)
14. Can we centralize administrative tasks at the National Processing Center?
15. Can we extend the temporary background check for survey staffs from 6 months to 1 year?
16. How can we improve the process CIPPS-based for decennial census promotions?
17. Can we hire Decennial Census staff virtually?
18. How can improve the process when we must implement retroactive terminations for field employees?

Project: Census Integrated Personnel and Payroll System

CIPPS aims to move our existing Schedule A temporary benefit employee personnel and payroll system to an Office of Management and Budget or Office of Personnel Management (OMB/OPM) approved Human Resources Line of Business Shared Service Centers (HR LoB SSC) provider and or commercial service provider.

Research Topics

- Recruiting, selection, and hiring.
- Background Checks.
- Online Training.
- Fingerprinting and Badging.
- Pay and benefits.

Research Questions

1. How to move our existing Schedule A temporary benefit employee personnel and payroll system to an OMB/OPM approved HR LoB SSC?
2. Will we need to renew our contracts with DAPPS?
3. How to improve our current manual and automated processes for our integrated approach to recruiting, tracking, hiring, and onboarding Schedule A vacancies in a timely manner?
4. What existing manual processes can be automated?

Project: Work-at-Home Staffing Model for Respondent Contact Center Operations

This project determines the feasibility of utilizing a Work-at-Home workforce for respondent contact center operations to add flexibility in meeting requirements and improving the caller experience during production. Working at home has many benefits like cost savings, desirability for employees, and the ability to capitalize on a more skilled and diverse workforce (such as hiring specific language skill sets). The project builds on experience with the Census Questionnaire Assistance (CQA) operation during the 2020 Census where a Work-at-Home program was implemented on the fly using hired and trained CQA contact center staff unable to work at their designated brick-and-mortar location due to COVID-19 enforced staffing limitations created by social distancing guidelines.

Research Topics

- Best practices for utilizing work-at-home staff for contact center operations.
- Security, technology, and equipment for work-at-home staff.
- Process difference for work-at-home staff (in comparison to staff located in physical contact centers).
- Communication mechanisms for work-at-home staff.
- Compensation model and cost elements for work-at-home staff.

Research Questions

General

1. How do contact centers associated with other Government agencies utilize a work-at-home model for staff providing services to the public?
2. What aspects of commercial contact centers utilizing a work-at-home model might be applicable to a census respondent contact center work-at-home design?

Security, technology, and equipment

3. Can the operational agility of respondent contact centers be improved by employing Work-at-Home staff? (E.g., can work-at-home staff be used to more efficiently and more effectively manage periods of unexpectedly high call volume?)
4. What is the approval process from Security in order to employ a work-at-home model for contact center staff?
5. What is the ideal working environment needed to protect the confidentiality of data, especially Title 13 data, for work-at-home staff?
6. What equipment would work-at-home staff use? (Options include “Bring Your Own Device,” “Government Furnished Equipment,” or “Vendor Issued.”)
7. How would processes and procedures be enforced in a work-at-home environment? Are there processes and procedures in place as a part of the Field program that could inform a work-at-home program?
8. What is needed to ensure adequate sound quality for work-at-home staff when on respondent calls, in terms of background noise and voice quality?
9. How would the enterprise monitor internet quality and connectivity such as clarity, bandwidth and latency for Work-at-Home staff?

(continued on next page)

(Continued)

Project: Work-at-Home Staffing Model for Respondent Contact Center Operations

Processes

10. What would be the differences in processes, procedures, and technology for Work-at-Home staff versus physical location staff?
11. How would equipment be distributed to Work-at-Home staff upon being hired?
12. How would technical, hardware, and software issues be resolved for Work-at-Home staff?
13. How would training be conducted for Work-at-Home staff?
14. How would quality monitoring be conducted for Work-at-Home staff?
15. How would equipment be collected from Work-at-Home staff upon termination or job abandonment?

Communication

16. How would Work-at-Home staff escalate issues to management while on a live call?
17. How would workforce management and supervisors communicate with Work-at-Home staff?

Other

18. What contact center positions would be eligible for a Work-at-Home program?
19. Are Work-at-Home staff limited in the type of calls they can handle? (Example might be data collection vs. assistance calls)
20. How would the compensation model for Work-at-Home staff differ from physical location staff? Would internet access be reimbursable for Work-at-Home staff?
21. What cost elements would be affected by utilizing Work-at-Home staff when compared to a traditional design where all staff are located in physical contact centers?

EA 4: Streamline Operational Support Infrastructure

Project: Internal Partnership Training and System Support

This project includes three sub-projects that seek to apply several lessons learned from the 2020 Census Partnership Program to improve this program for the 2030 Census. These improvements aim to reduce cost, increase quality, and increase operational efficiency. Assessments and "lessons learned" studies of the Integrated Partnership and Communications (IPC) program identified several areas of improvement for the 2030 Census Partnership Program. In these studies, improvements were recommended for three areas: (a) the Customer Relationship Management system (CRM), (b) partnership staff training, and (c) the internal distribution of census materials and promotional items to partnership staff.

Research Topics

- Improvements to the CRM database.
- Improvements to planning, processes, training and roles regarding the use of CRM.
- Redesign the partnership staff training plan.
- Improve the quality of partnership staff training by making it effective, consistent, timely, and up to date continuously.
- Redesign the process for distributing census materials and promotional items to partnership staff.
- Improve the 2030 Census processes to receive internal clearance for materials and to create and distribute them.
- Identify contingency plans should materials and items not arrive on time in the field.

Research Questions

1. How can the CRM be improved to be more user friendly, accommodate users of various technological experience, and be integrated with web analytics and internal systems for the 2030 Census?
2. How can partnership staff training be improved to make it more engaging and consistent in quality across all training topics?
3. How can the internal distribution of 2020 Census promotional materials and promotional items to partnership staff be improved to better ensure that (a) they are delivered on time and (b) there is a plan in case they arrive late?
4. If materials are obtained from a contract (as they were during the 2020 Census from the Integrated Communications Contract), what material distribution requirements should be accounted for in the contract?

EA 4: Streamline Operational Support Infrastructure

Project: System Information Hub (SIH)

This project will develop the System Information Hub (SIH), which will provide a one-stop destination for Decennial Census staff on system related information. This site will provide general (system owner, development lead, environments available, Authority to Operate (ATO) dates, etc.) and specific system information (system description, system integration, hosting location, whether the system is being developed to support an enterprise area or only the 2030 Census, or only for on-going survey support). In addition, the SIH will display a filtered list of systems being used to support different functional areas, operations, and provide internal links to specific system SharePoint sites.

The 2020 Census had around 80 different systems to complete the decennial census. The management of the systems became a challenge to keep up with for the average headquarter staff with the following needs: identifying the most appropriate system to system integration details and researching the ability to automate and display the system information. This project will first develop a prototype for the 2020 Special Census. If a successful pilot research phase is obtained, the SIH may be recommended for further 2030 Census testing and the 2030 Census.

Research Topics

- Develop a SIH prototype.
- Automate a process to efficiently display general and specific information that has a high level of content change.
- Develop a central place for stakeholders to find information for operations and the systems used to accomplish their daily tasks.
- Improve research efficiencies and reduce training time for new staff.

Research Questions

1. How can we integrate general and specific system information into a single accessible location?
2. How can system information that currently exists in various forms and locations across the Bureau, be accessed and displayed in a single location, site, or portal?
3. How can this information be maintained and automatically kept up to date without increasing burden on system owners and development teams?

EA 5: Continuous Data Collection and Aggregation

Project: Administrative Data Quality and Usage

This project will conduct administrative data research in order to provide the ability to produce linked location and demographic data at any given instance throughout the decade to allow continuous assessment of coverage and quality. This project will do this by 1) assessing the current inventory of administrative data to identify areas of improvement; 2) identifying and acquiring new, trusted administrative data sources for 2030 Census operations to meet these areas of improvement, such as state-level data sources like vital records; 3) identifying reliable data sources for operations that did not use administrative records (AR) in the 2020 Census, such as in-office enumeration in Island Areas; and 4) creating, managing, and assessing the fitness for use of the creation of, and continuous updates to, the 2030 Census Person Characteristic Frame (PCF) as a whole.

Research Topics

- Assess the 2020 Census administrative data composite.
- Acquire, assess, and integrate new administrative data sources.
- Identify and assess administrative data sources for new uses.
- Create and assess the 2030 Census PCF.

Research Questions

1. What are the strengths and weaknesses of the 2020 Census AR composite with respect to covering the U.S. population?
2. How does the 2020 Census vintage of the administrative data inventory compare to self-responses and household interview responses from the 2020 Census? Comparisons include demographic data, address population counts, and housing unit status.
3. Are there populations, demographic groups, or geographies for which AR coverage could be improved?
4. Are there additional data sources that improve the person and characteristic coverage and utility of the 2020 Census composite inventory, including those currently accessed by other areas of the Census Bureau?
5. Which AR sources help improve coverage of historically undercounted populations?
6. Can vacant housing units be identified using new data sources such as utility data, cell phone data, broadband internet data, and real estate and rental websites? Is it feasible to obtain these data sources?
7. Is it feasible to acquire and use ARs to enumerate select group quarters?

EA 5: Continuous Data Collection and Aggregation

Project: Decision Rules for Effort Allocation Leveraging Information from Administrative Data Sources

This project will investigate methods to reduce field effort and respondent burden during census data collection operations ([including In-Field Enumeration), for the U.S. household population by assessing whether, and under what conditions, administrative data could provide the information. All new methods will be simulated and tested to understand impacts to data quality. This project will provide rules about how to use administrative data to allocate data collection between in-office enumeration and field enumeration in order to reduce respondent burden and in-field workloads for census data collection to a greater extent than for the 2020 Census.

Research Topics

- Develop and evaluate decision rules for In-Office Enumeration.
- Test decision rules on simulated and historical census data.
- Assess the impact of decision rules.

Research Questions

1. What are appropriate frameworks to consider using information from administrative data in the 2030 Census? Examples could include minimizing burden and data missingness rate; or minimizing data missingness rate and data collection attempts.
2. How can past and current research into predictive models for quality, progress, burden, and resource usage be adapted to the identified frameworks?
3. Can predictive models necessary for the identified frameworks be constructed for 2030 Census research?
4. What are the predicted impacts on quality, progress, burden, and resources when using administrative data according to the defined frameworks?

EA 5: Continuous Data Collection and Aggregation

Project: Research on the Undercount of Young Children Using Administrative Data

This project will use administrative records (AR) matched to decennial census data to investigate the undercount of young children. These data will allow the characteristics of children who were not counted in the 2020 Census as well as the characteristics of households with young children who were not counted to be described. In addition, AR matched to 2020 Census data will allow the accuracy of the information provided in the census, if any, to be determined. This approach will also allow the project team to explore whether children were missed in the decennial count because they were in a household that was completely unenumerated or because they were specifically omitted from the roster in an enumerated household. The existing literature on the undercount of young children has highlighted this distinction between unenumerated households and unenumerated children in enumerated households: households in multi-unit buildings with difficult access or low-income multi-generational households are examples of hard-to-reach households, while children with more than one usual residence or in a temporary living arrangement are more likely to be unenumerated.

Research Topics

- Document the undercount of young children in 2020 Census.
- Identify households with children at risk of undercounting.
- Determine implementation of undercount methodology for 2030 Census.

Research Questions

1. Can the Census Bureau use AR to understand the child population that was undercounted in the 2020 Census?
2. What AR data sources are useful for measuring and predicting undercount?
3. What additional sources of AR data can be identified to improve the accuracy and coverage of young children?
4. What was the extent of the undercount of young children in the 2020 Census?
5. What are the demographic characteristics of undercounted children in the 2020 Census?
6. What are the characteristics of households with undercounted children in the 2020 Census and are these enumerated households omitting young children or is the entire household missing from the census count?
7. What are the characteristics of neighborhoods – measured at the census-tract level – where children are at higher-than-average risk of being undercounted?
8. Can the Census Bureau develop a predictive model using AR to identify households that are likely to have children and their likelihood or risk of experiencing undercount?
9. To what extent can modeled output guide the use of administrative data to directly add young children to a census response and thus increase accuracy?

Project: Improving Within-Household Coverage Using Administrative Data

This project will explore methods for using administrative data to either: add people into households where the census response omitted the person or remove people from households where the census response inappropriately included them. The primary interest in adding missing people is to correct the persistent undercount of young children and historically undercounted populations, although all ages and populations will be explored. The project will seek to determine the accuracy of these added people and removed people by means of a follow-up interview during one or several mid-decade tests.

Research Topics

- Demonstrate using administrative data to reduce under- and overcount.
- Improve household coverage using administrative data.

Research Questions

1. Can administrative data be used to add people to census responses to improve the household count?
2. Which administrative data are appropriate for this use?
3. What rules or models should be used to add people to census responses? To delete people or move them to a different address?
4. Does this process increase duplication (overcoverage) in the census?
5. Does this process increase undercoverage?
6. How often are rosters modified due to the results of in-office enumeration efforts during processing?

EA 5: Continuous Data Collection and Aggregation

Project: Improving the Listing and Enumeration Address Filters

This project will study the address filter used to determine the housing unit, group quarters, and transitory location frames for the 2020 Census. The census listing filter is used to identify the best set of living quarters to validate or update in those geographic locations that require in-field listing. This may include some addresses for which there is conflicting information in the Master Address File (MAF) from various sources, including the Delivery Sequence File (DSF) from the United States Post Office. The census enumeration address filter determines the set of living quarters to send forward to the enumeration operations and incorporates information obtained during listing operations. This research aims to identify attributes of MAF addresses that can be used in the address filters to help identify living quarters that do (or do not) belong in a listing or enumeration frame. Objectives include identifying attributes of existing MAF records that are excluded from the address filters erroneously; reducing duplication in the filtered address list, which leads to over-coverage in the address frame; and identifying attributes of MAF records included in the address filters erroneously.

Research Topics

- Addresses not selected from MAF but responded anyway.
- Addresses selected from MAF but not ultimately enumerated.
- Centrally located community mailbox addresses.
- American Community Survey and decennial filter differences.
- MAF records enumerated in 2010 but later dropped off the DSF.

Research Questions

1. How does the set of addresses identified by the census address filter for the 2020 Census compare to the final set of addresses counted in the census?
2. How can the address filters be revised to provide better coverage?
3. Do the results of 2020 Census response person-matching identify potential address duplication in the address list of filtered records?
4. How can “undeliverable as addressed” data be used in the census address filters?

EA 5: Continuous Data Collection and Aggregation

Project: Person and Address Matching Research

A number of systems in the 2020 Census used statistical matching applications to perform record linkage of person or address records for various purposes. For some purposes, this record linkage attempted to match census responses to other sources of data (such as administrative data). For others, the census responses were linked to themselves in an attempt to identify possible duplication. The systems that used record linkage include (but are not limited to) Self-Response Quality Assurance (SRQA), Production Environment for Administrative Records Staging, Implementation and Storage (PEARSIS), Decennial Response Processing System (DRPS), and Sampling, Matching, Review and Coding System (SMaRCS). Additionally, each of these systems had different requirements for when two (or more) records would be considered a match (or a linked pair), thus any software used had to accommodate these different requirements. This project will research alternative matching software for the 2030 Census, test implementation of that software on 2020 Census data, and compare the results to those obtained during 2020 Census production processing.

Research Topics

- Compare matching software to previously used methods.
- Identify best matching software for 2030 Census.

Research Questions

1. How do the alternative systems and software results compare to those obtained by 2020 Census processes?
2. Can a single alternative system or software handle the variety of matching requirements needed by different operations?